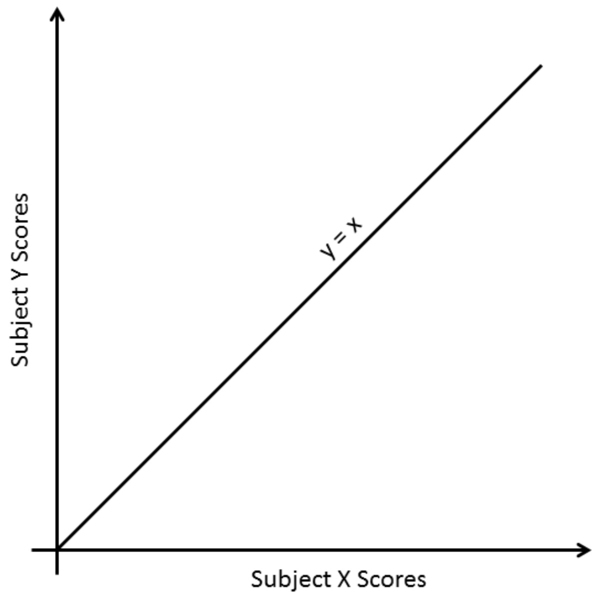
Theory: A Possible Testing Method

Assuming that a cohort of students find two courses:-

- to be equally difficulty.
- are assessed to an equal standard.
- are taught to the same standard.

then you would expect that the results from these two courses to form a straight line with a=0 and b=1, i.e. the students would get exactly the same results in both.



The actual regression line will be different - and that needs assessing before the introduction of any new materials or methods. The data for as many previous years as possible should be examined across as many courses as possible. A data group could be formed for each course combination of those students who completed both. The correlation and regression line could then be calculated for each course pairing for each academic year.

Changes in teaching method or changes in teaching staff would show up as changes to the slope, *b*, between years since student preference for the two courses could be altered. However, a complicating factor would be that the introduction of new materials or teaching methods could also alter student preferences. A longitudinal study would be expected to show up each of these factors as changes between year groups which are static and those which are dynamic are examined.

Student perception aside, the introduction of new materials would be expected to keep the slope of the line the same (the brighter and more dedicated will continue to achieve better than those less able and less motivated). Therefore the intercept could be expected to show the effect.

If the changes to *a* are statistically significant then the new materials would have had an effect. The degree of that effect (or not) could be shown by its distance from the expected value of *a*.

Another measure could be the effect on the correlation coefficient. In an ideal world, it would be found that $r = 1$. For that to happen all the students would have to achieve the same normalised score (taking the regression line into account) for each course – and that would be suspicious since students vary.

These differences in performance for students between units (learning dissonance) should diminish if the new materials or teaching methods an effect. It will not affect those below the regression line as much since these students are performing better in comparison course. However, those who have more difficulty with the altered course will have access to new materials and/or alternative learning modes and would therefore be expected to improve. The degree of improvement in $r^2 \times 100$ (or not) would be another test of the effectiveness of changes in the teaching environment.

Finally, the degree of learning dissonance between courses:-

$$\sum \sqrt{(y - y_{exp})^2}$$

would be expected to reduce as standard of teaching becomes more uniform. Statistically large discrepancies (a term that would need to be quantified) might be an area of concern.

Practice: The 2010/11 academic year results

The tables in the next section show:-

Table 1 : The number of students who took both the course. The numbers are not the same for each cell since students from other programmes may also have participated.

Table 2 : The Pearson Product Moment Correlation Coefficient (r) between the two courses for the scores in each of these courses. Any scores starting with L were deemed to be 40.

Tables 3 and 4 : The intercept and slope for the linear regression line between these two courses.

Table 5 : The theory outlined in the previous section hypothesises that the relationship between these scores would be linear. This has been tested by sorting the scores for the y- and x-variables and then using linear relationship to estimate the y-variable. The Durbin-Watson d-statistic should be near to 2 if the relationship is linear. Ten of the results (indicated with *) show a significant level of autocorrelation which indicates a non-linear relationship. This has not been explored further yet.

The row of interest is the top one for COMP1148, since that is the course on which the proposed system will eventually be trialled.

If the new system were to have an effect then the correlation coefficient would be expected to improve because the results would bunch more at one end. If both the mean score ($\bar{x}$) and the correlation coefficient (r) increase then the effect would have been beneficial.

Mutual correlation could be a problem for this approach. If COMP1148 becomes more understandable then students may feel generally more motivated and achieve better scores in other units too. The degree of effect can be estimated using the changes in *r* relative to COMP1148.

The only change to COMP1148 this academic year was the replacement of Kate Finney. The materials were the same and the method of marking was the same. It will therefore be an interesting exercise to see whether this one change made any difference that detectable statistically. This can be tested once first year scores have been finalised in September 2013. This test will show whether the statistical tool described here has any potential.

Potential further interesting areas

1. Could this approach be used to assess the effectiveness of new teaching materials as well as approaches to teaching?
2. Is there anything which can be learned about the level and effectiveness of other courses? Could it, for example, highlight areas of excellence or areas for improvement?
3. Could this approach be used in situations where non-numeric results are achieved, for example, fail, pass, merit and distinction as on National Diplomas? If so, one would have to use Spearman's Rank Correlation Coefficient instead.

Cross-Course Statistics for 2010/11 for those taking courses on the year 1 of the Computing programme

Table 1: n (Sample size - the number taking both courses)

	COMP1152	COMP1587	COMP1588	COMP1589	MATH1110	MATH1111
COMP1148	260	337	152	189	193	169
COMP1152		250	141	107	176	152
COMP1587			148	183	189	164
COMP1588				0	136	137
COMP1589					55	30
MATH1110						164

Table 2: r (Pearson's Correlation Coefficient

	COMP1152	COMP1587	COMP1588	COMP1589	MATH1110	MATH1111
COMP1148	0.8484	0.5946	0.7911	0.6673	0.6286	0.8004
COMP1152		0.6243	0.8252	0.6835	0.6578	0.8093
COMP1587			0.7046	0.3966	0.6741	0.6203
COMP1588				******	0.7156	0.8502
COMP1589					0.3333	0.4435
MATH1110						0.8146

Table 3: a (The intercept)

	COMP1152	COMP1587	COMP1588	COMP1589	MATH1110	MATH1111
COMP1148	16.014	36.295	10.460	30.652	31.353	10.057
COMP1152		29.412	0.090	24.484	20.785	-1.600
COMP1587			-2.544	42.100	19.897	7.610
COMP1588				******	27.796	5.754
COMP1589					35.957	32.595
MATH1110						-5.211

Table 4: b (The slope)

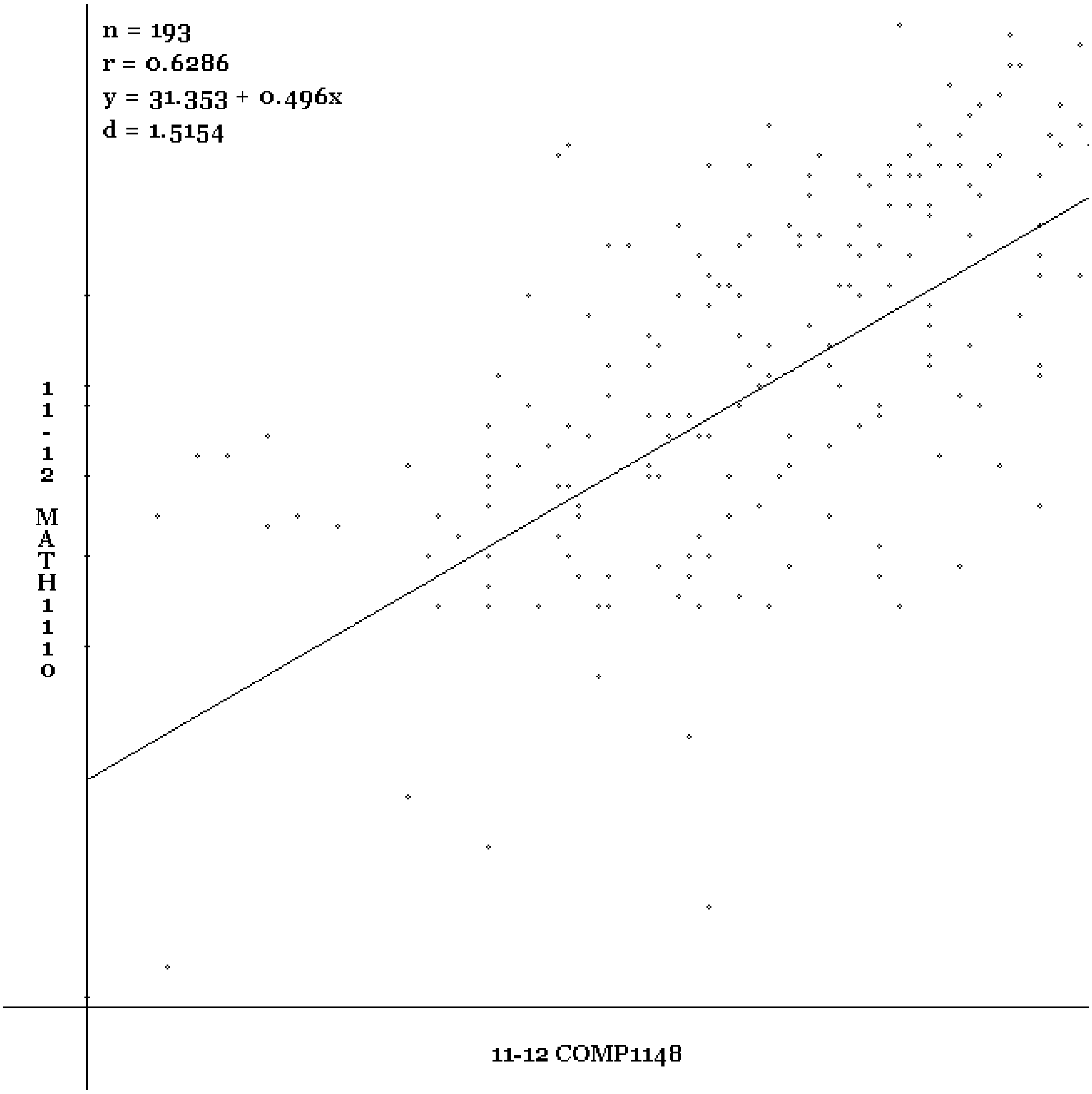
	COMP1152	COMP1587	COMP1588	COMP1589	MATH1110	MATH1111
COMP1148	0.685	0.514	0.713	0.610	0.496	0.770
COMP1152		0.681	0.914	0.780	0.686	0.980
COMP1587			0.816	0.387	0.614	0.715
COMP1588				*****	0.619	0.908
COMP1589					0.392	0.524
MATH1110						1.000

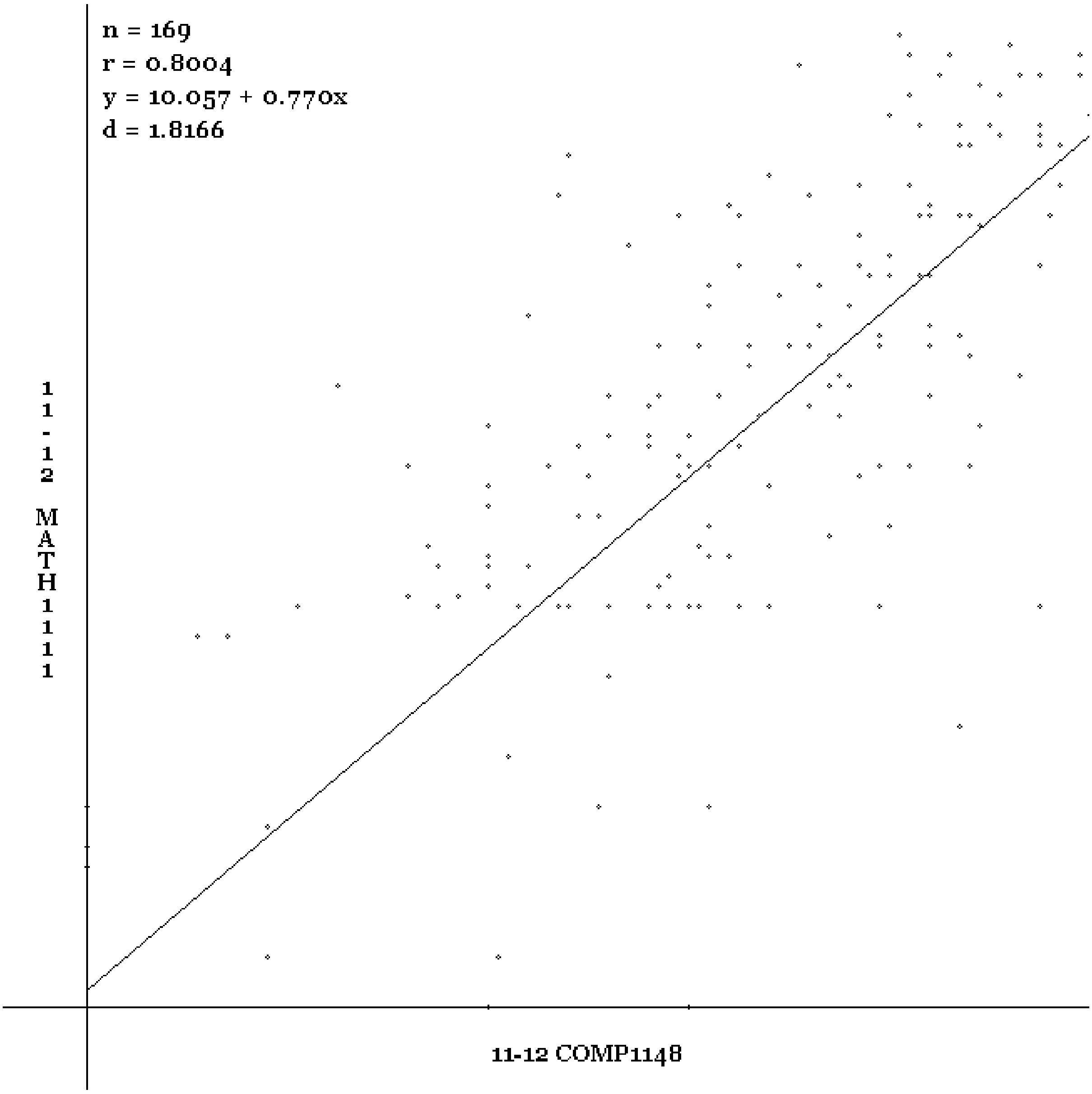
Table 5: d (Durbin-Watson d-Statistic)

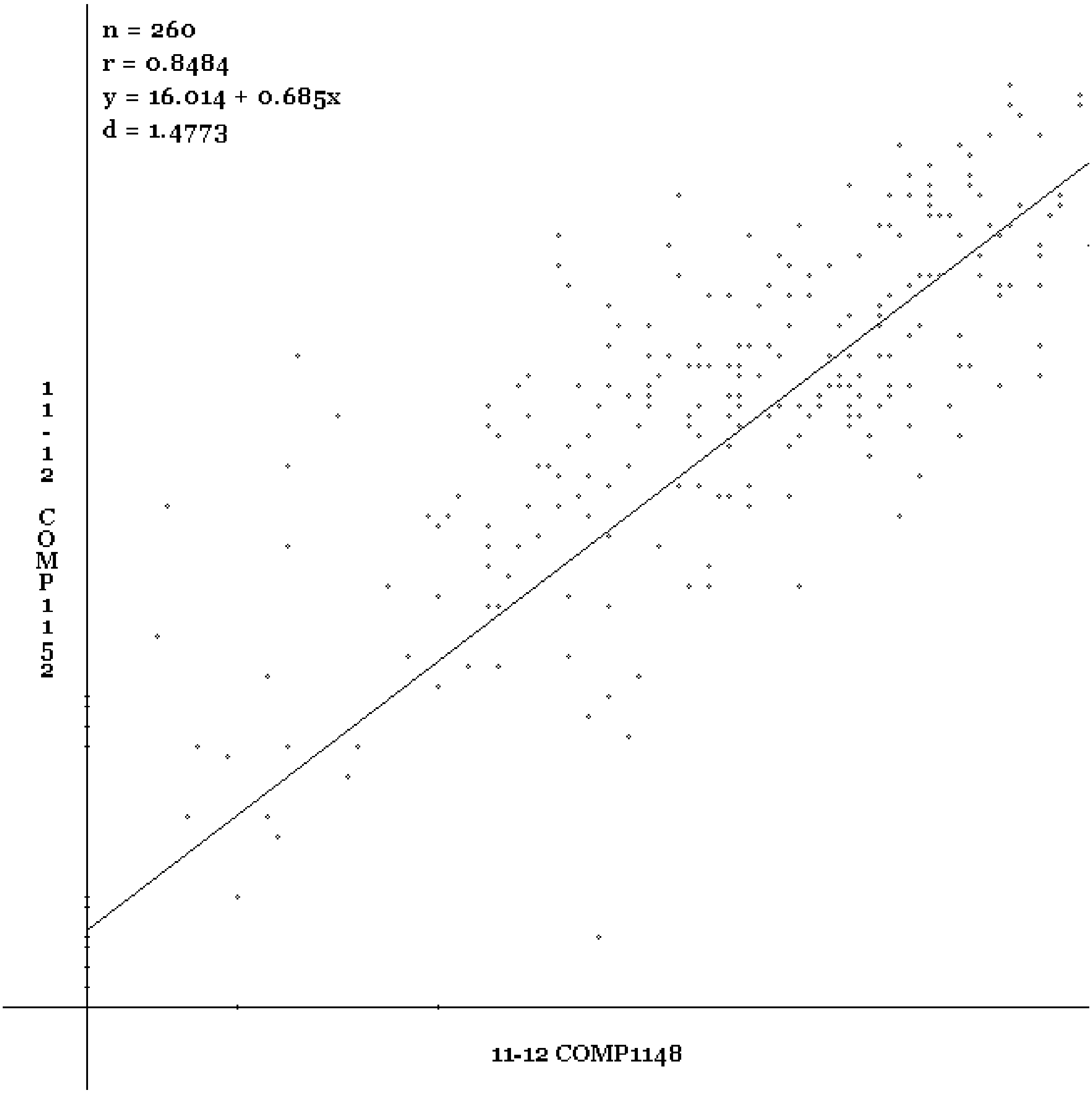
	COMP1152	COMP1587	COMP1588	COMP1589	MATH1110	MATH1111
COMP1148	1.4773	1.3615	2.0456	1.5980	1.5154	1.8166
COMP1152		1.7927	2.0437	2.1814	1.9259	2.0762
COMP1587			1.7225	1.5370	1.5538	1.6947
COMP1588				******	1.2925	1.5445
COMP1589					2.0089	2.0001
MATH1110						2.2303

Note: Some combinations (marked *****) are mutually exclusive.

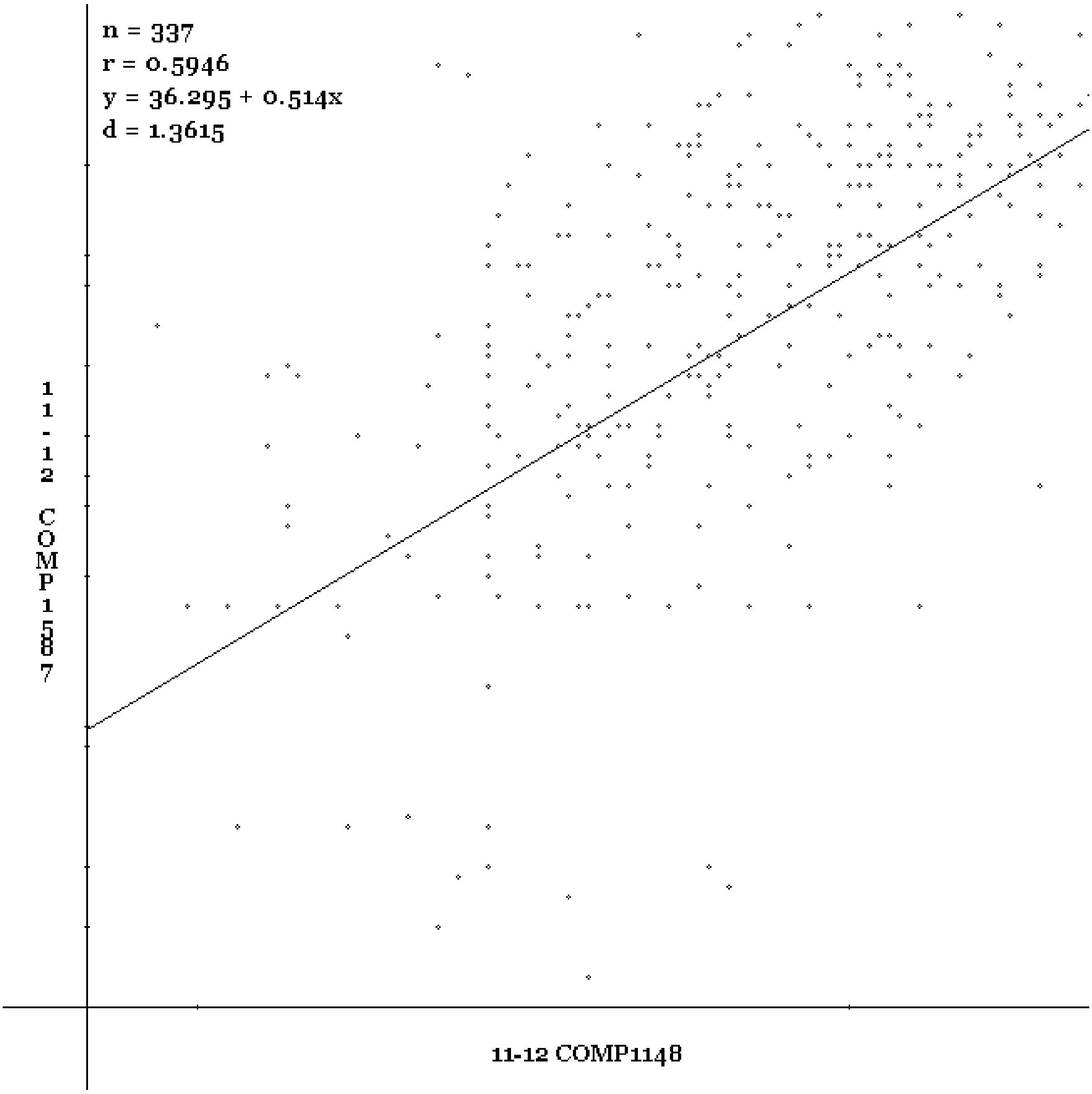
The Graphs

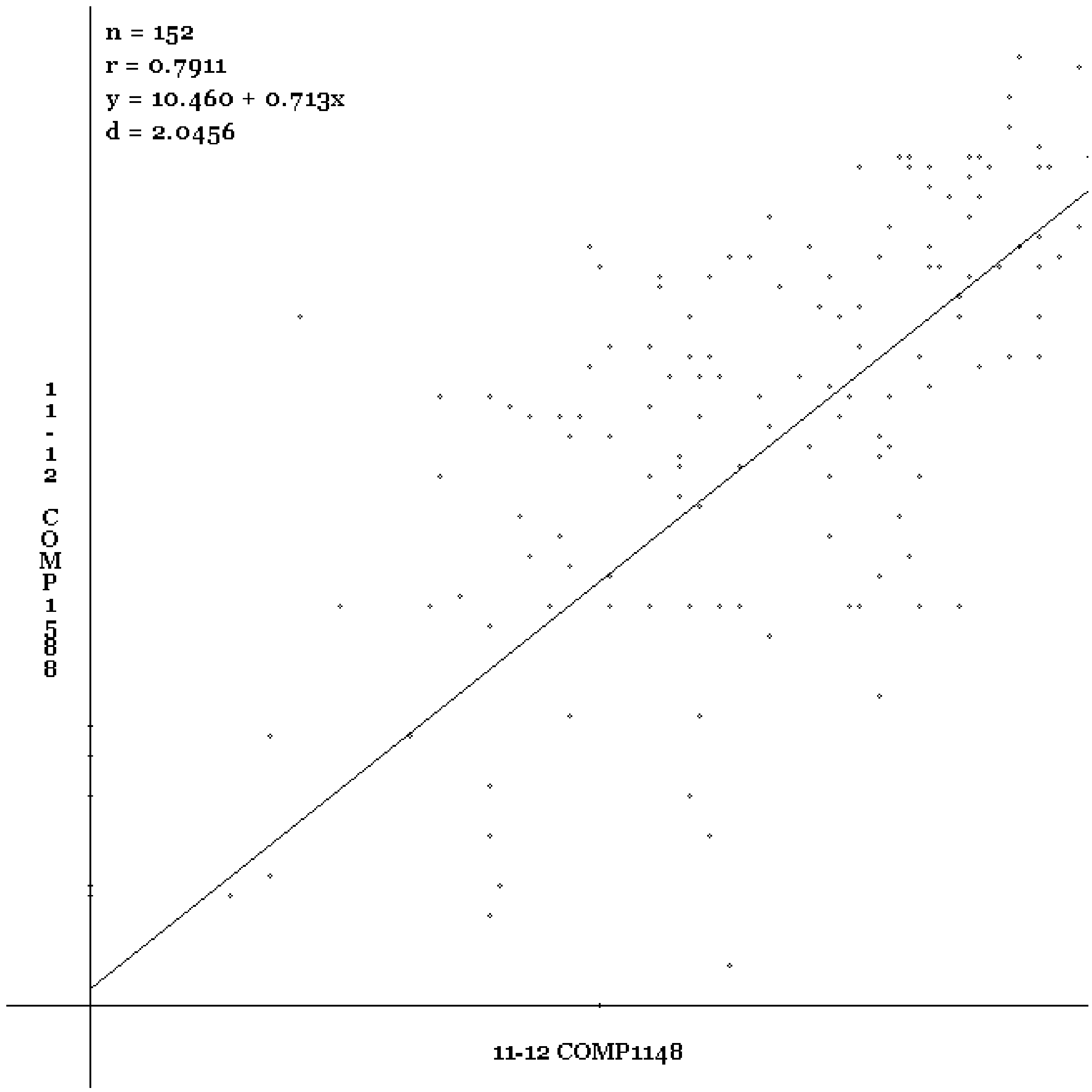

n = 193
r = 0.6286
y = 31.353 + 0.496x
d = 1.5154
11-12 MATH1110
11-12 COMP1148

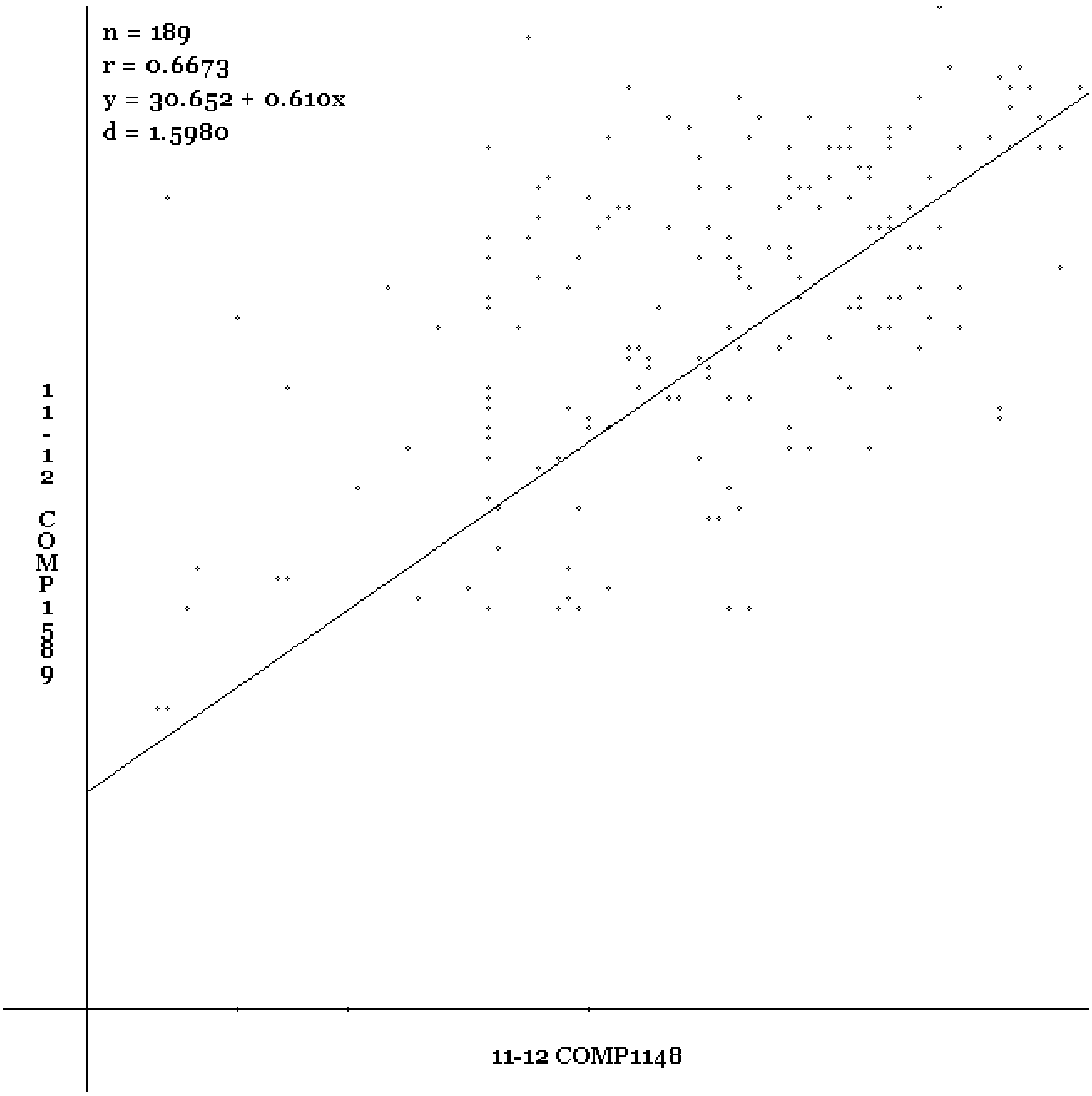

n = 169
r = 0.8004
y = 10.057 + 0.770x
d = 1.8166
11-12 MATH1111
11-12 COMP1148



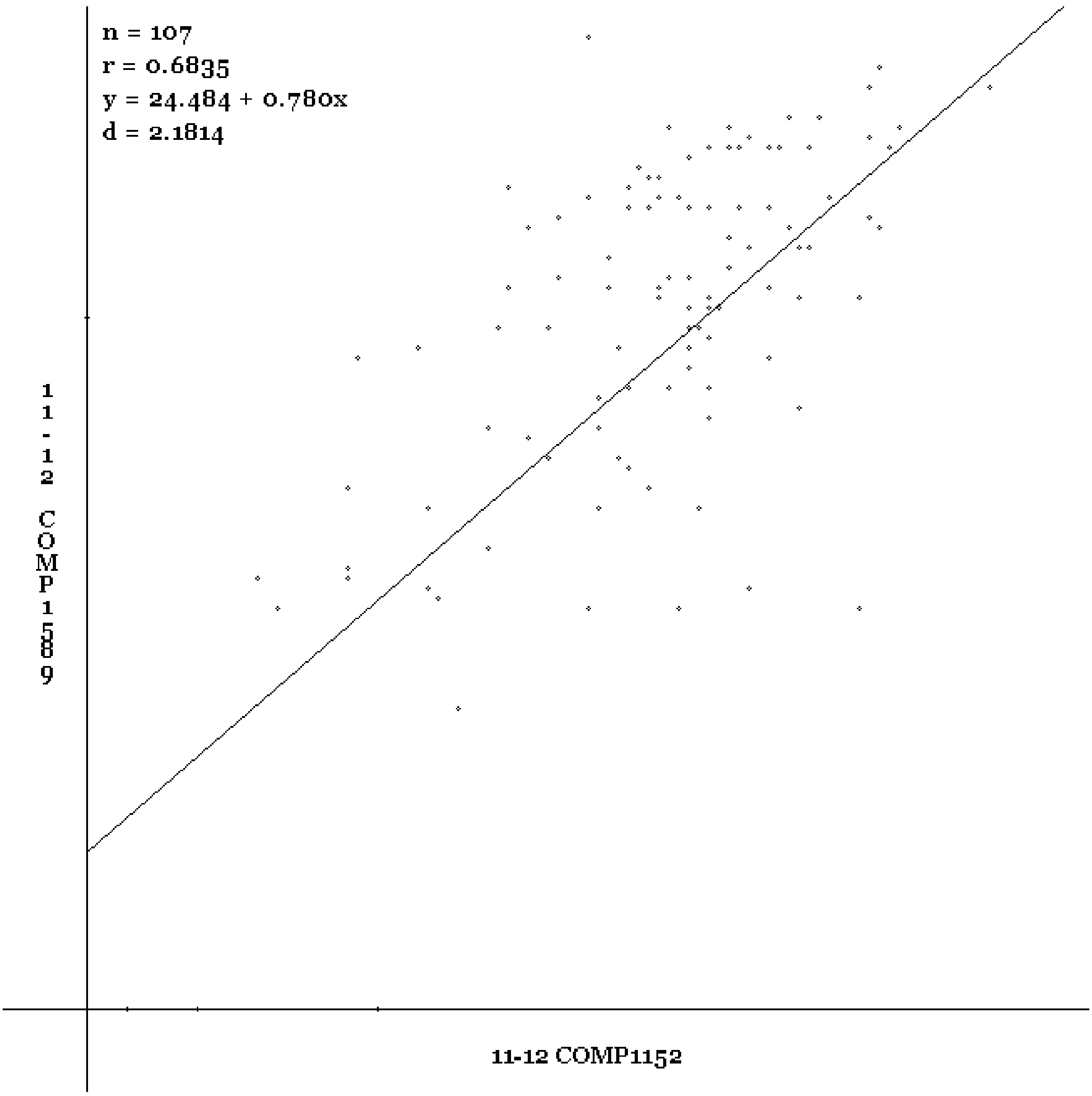
n = 260
r = 0.8484
y = 16.014 + 0.685x
d = 1.4773
11-12 COMP1152
11-12 COMP1148

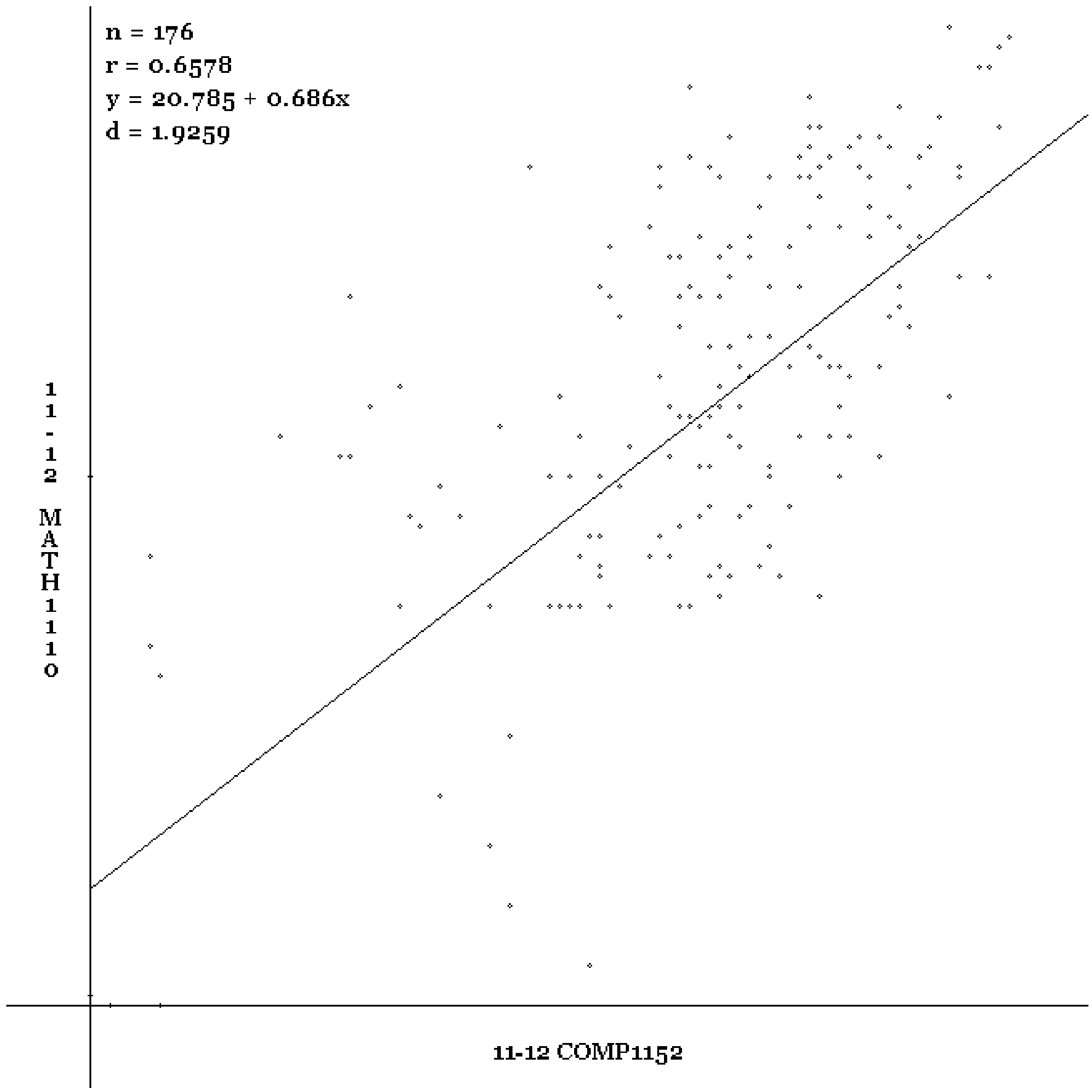

n = 337
r = 0.5946
y = 36.295 + 0.514x
d = 1.3615
11-12 COMP1587
11-12 COMP1148

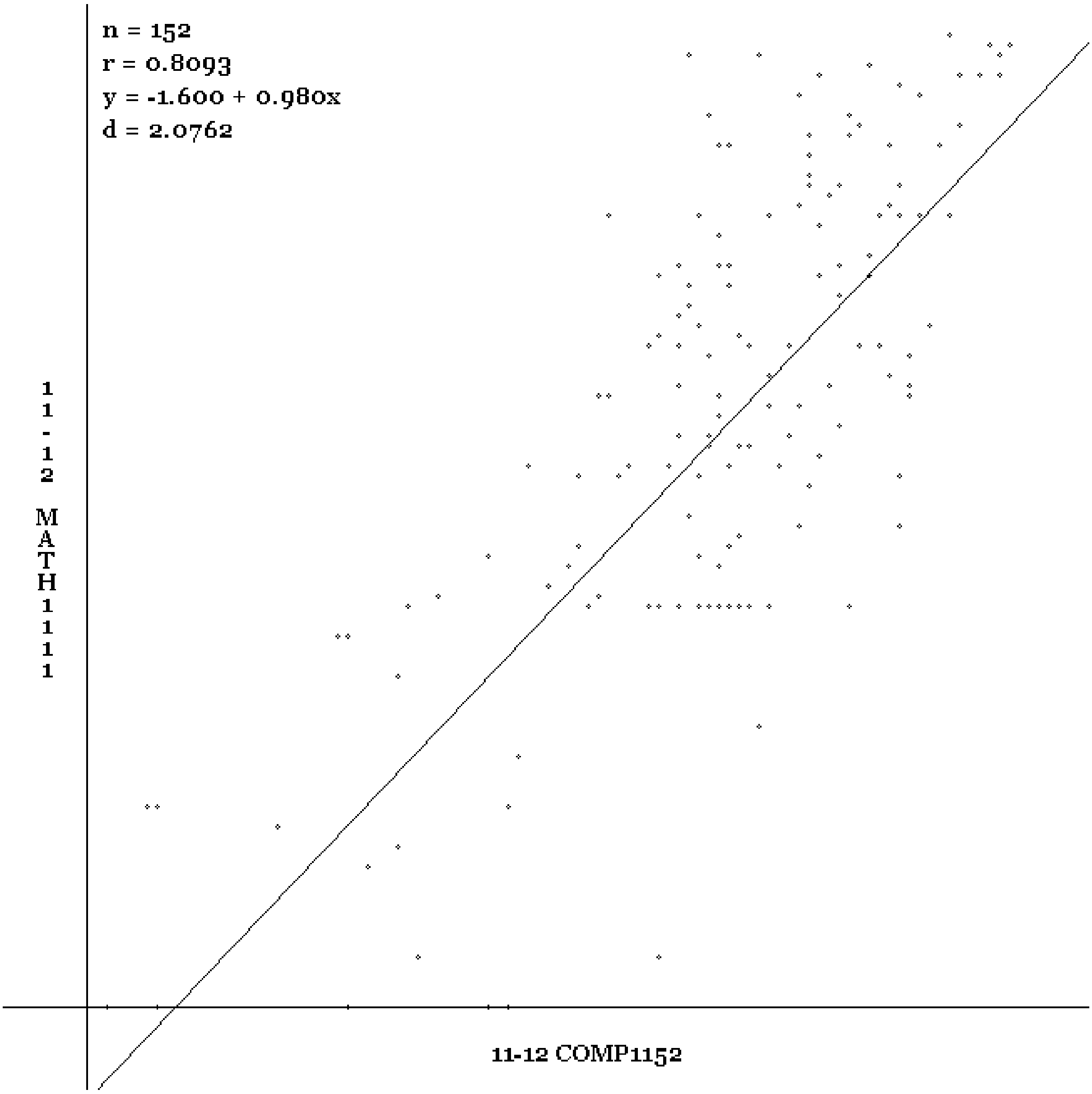

n = 152
r = 0.7911
y = 10.460 + 0.713x
d = 2.0456
11-12 COMP1588
11-12 COMP1148

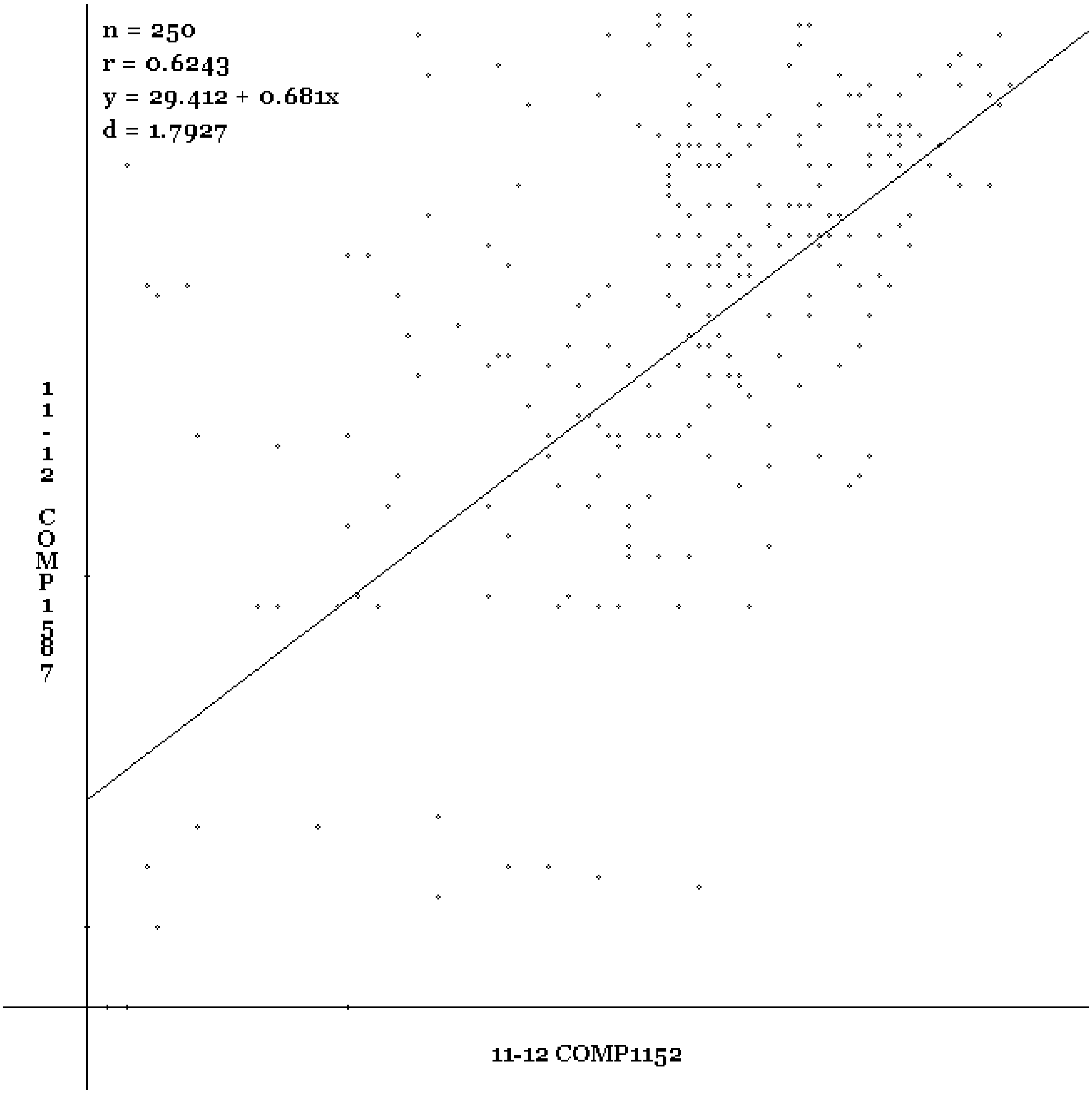


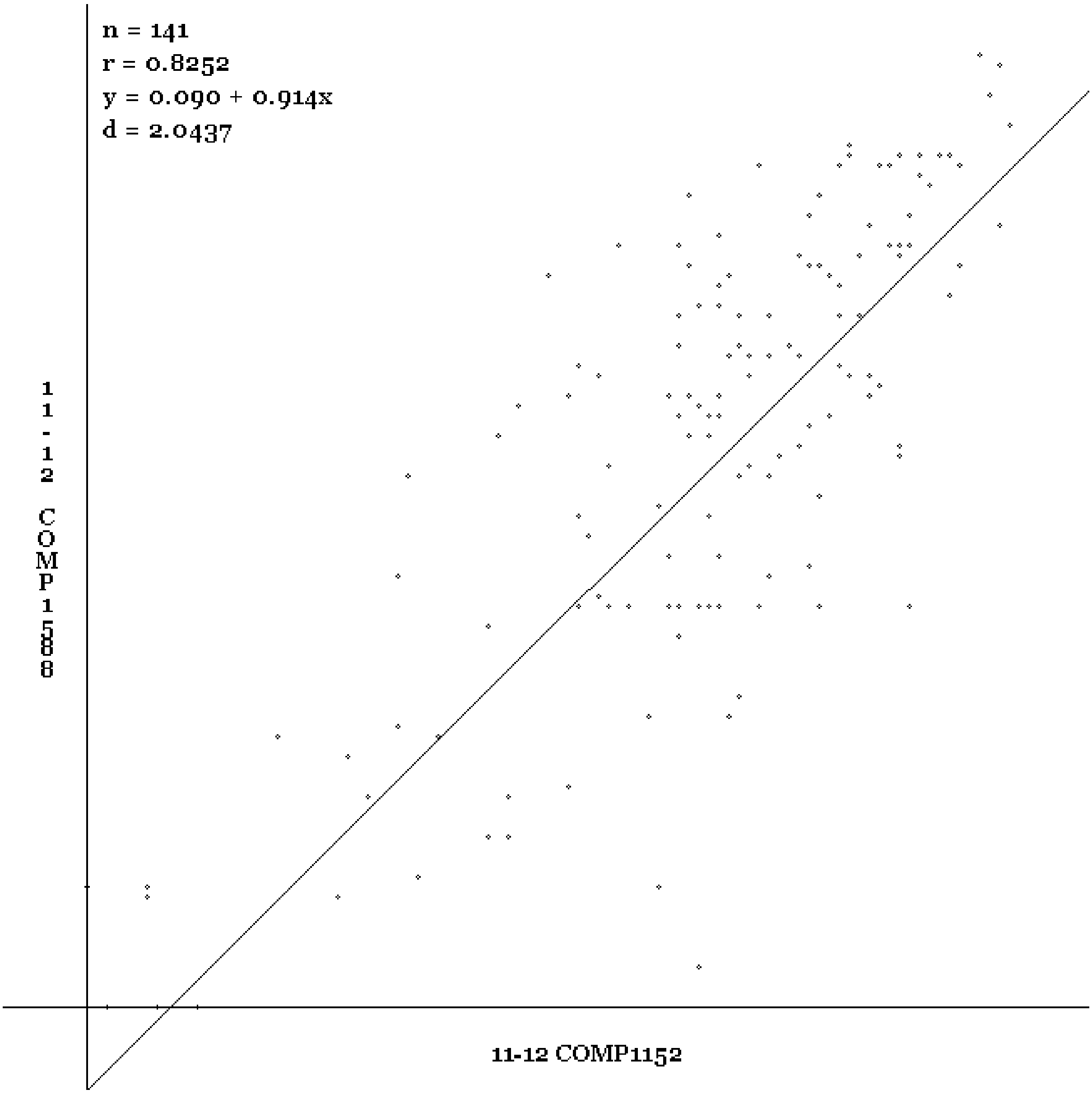
n = 189
r = 0.6673
y = 30.652 + 0.610x
d = 1.5980
11-12 COMP1589
11-12 COMP1148

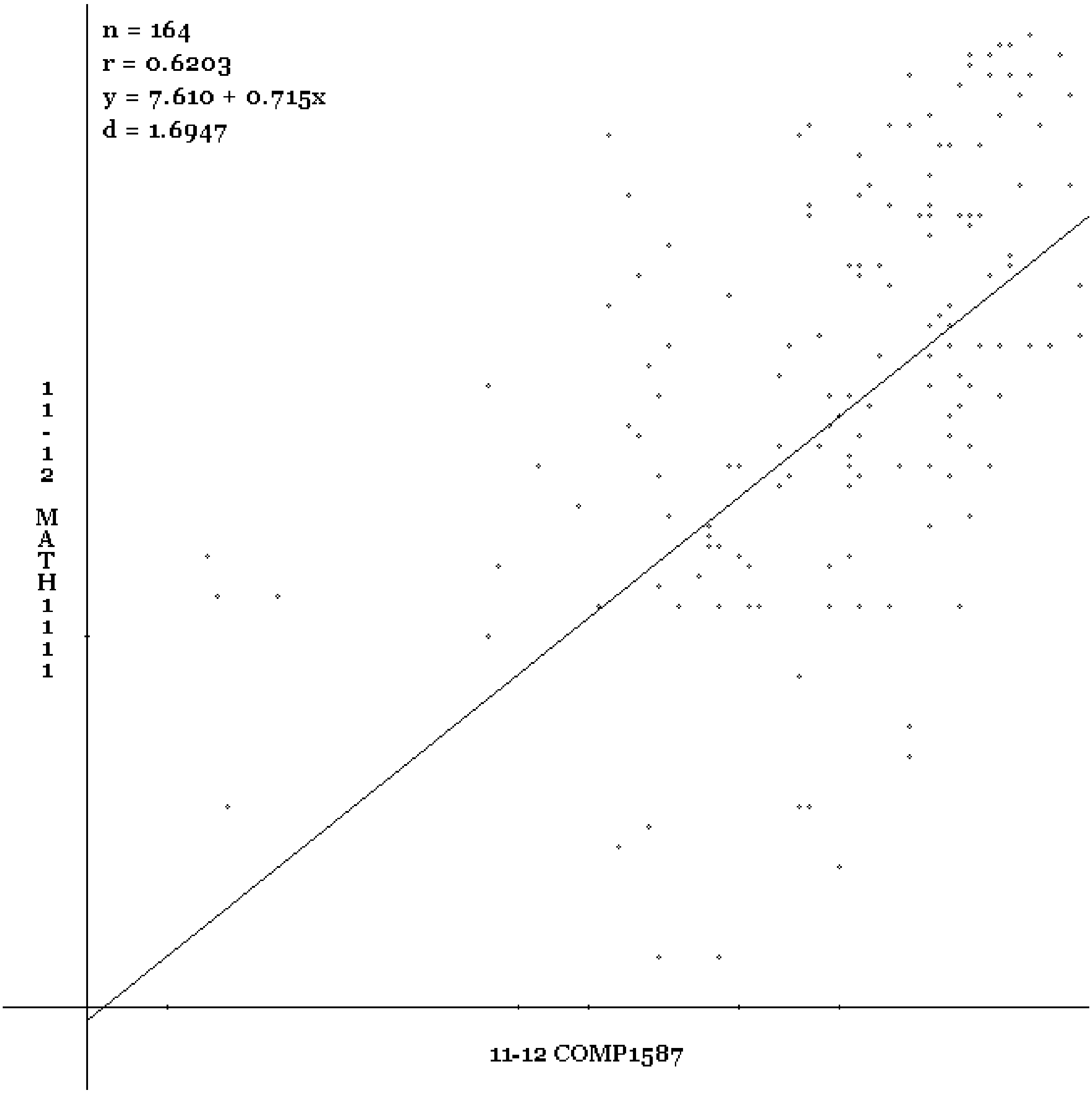

n = 107
r = 0.6835
y = 24.484 + 0.780x
d = 2.1814
11-12 COMP1589
11-12 COMP1152


n = 176
r = 0.6578
y = 20.785 + 0.686x
d = 1.9259
11-12 MATH1110
11-12 COMP1152

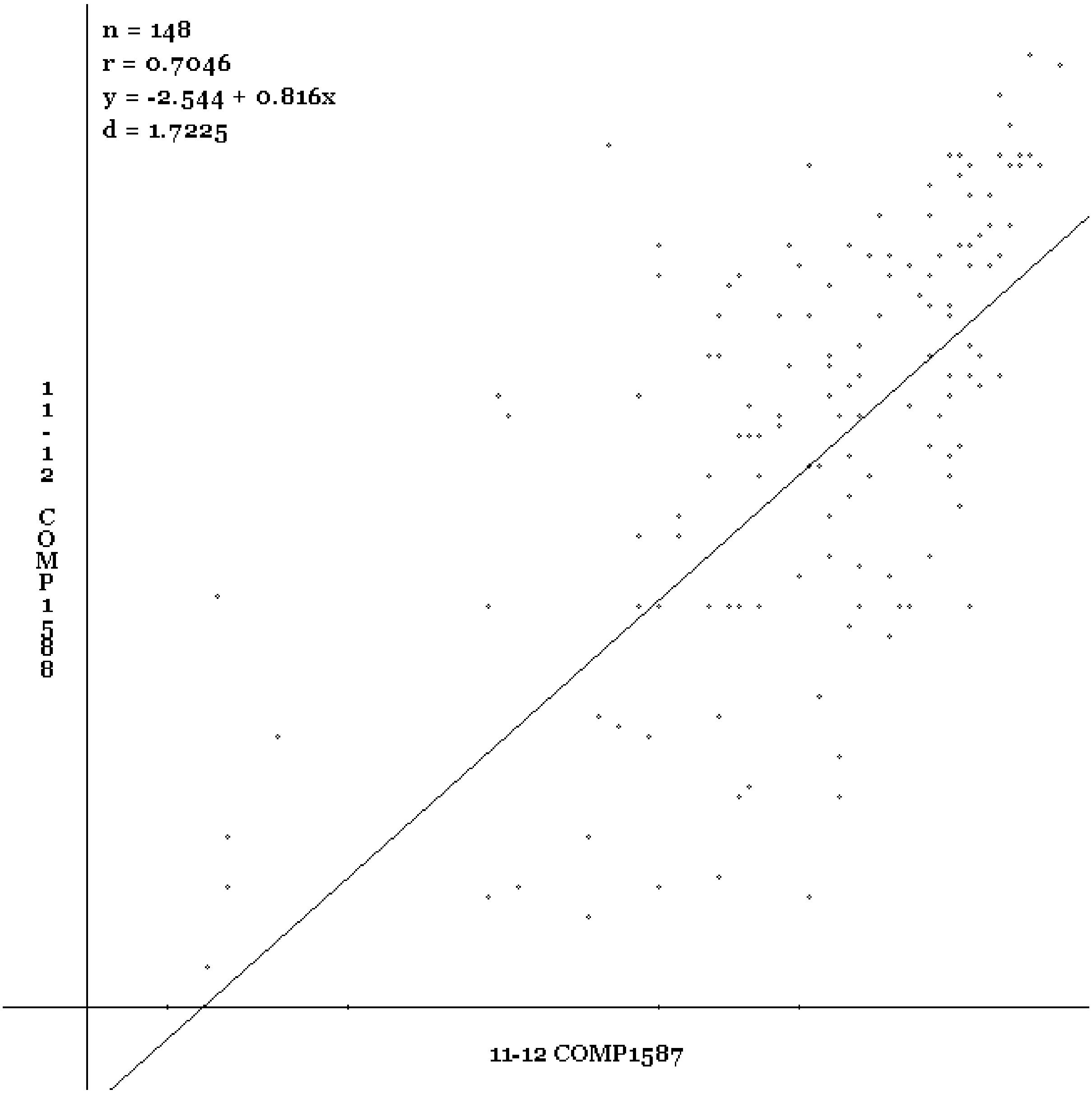

n = 152
r = 0.8093
y = -1.600 + 0.980x
d = 2.0762
11-12 MATH1111
11-12 COMP1152


n = 250
r = 0.6243
y = 29.412 + 0.681x
d = 1.7927
11-12 COMP1587
11-12 COMP1152

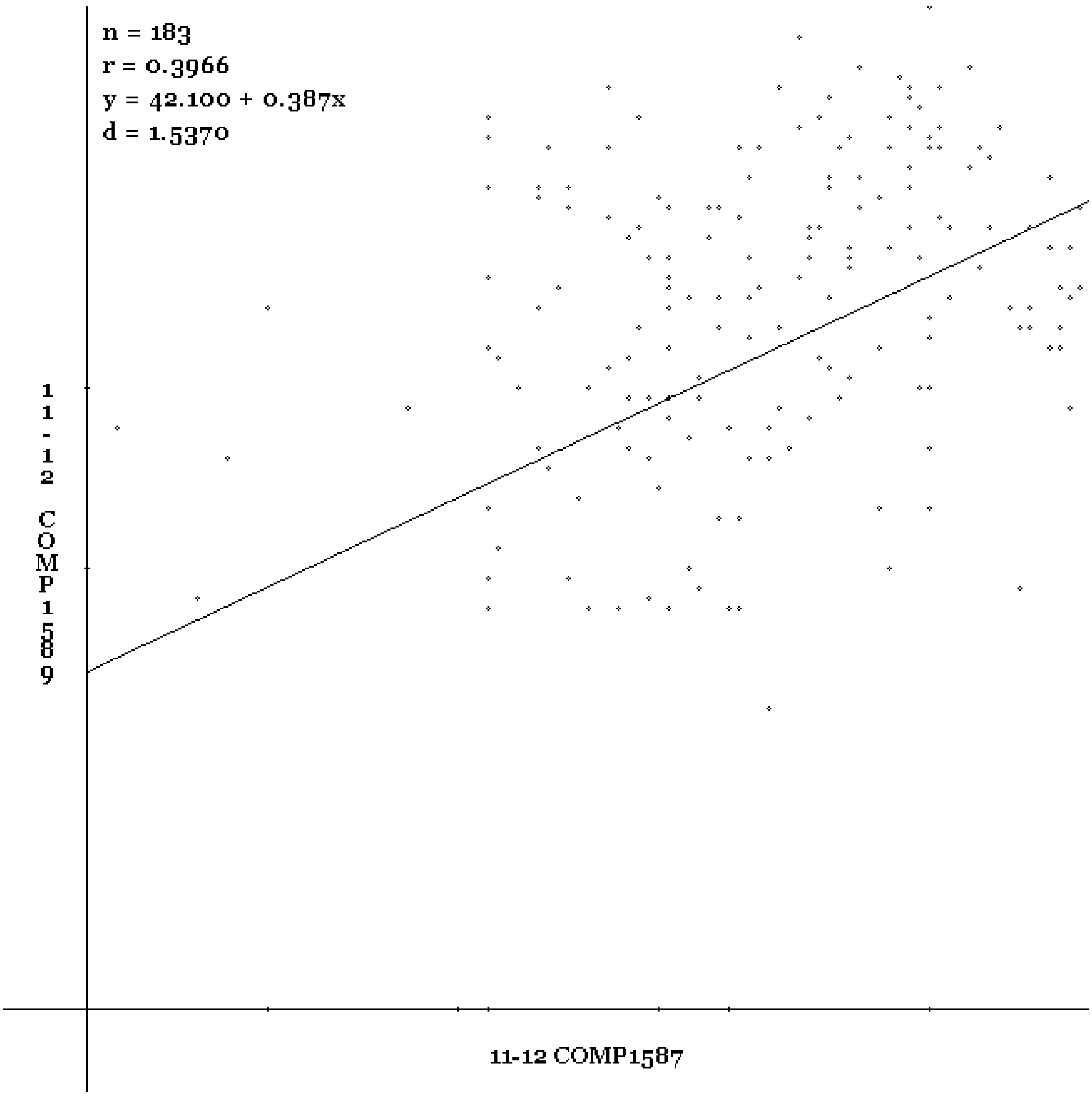

n = 141
r = 0.8252
y = 0.090 + 0.914x
d = 2.0437
11-12 COMP1588
11-12 COMP1152

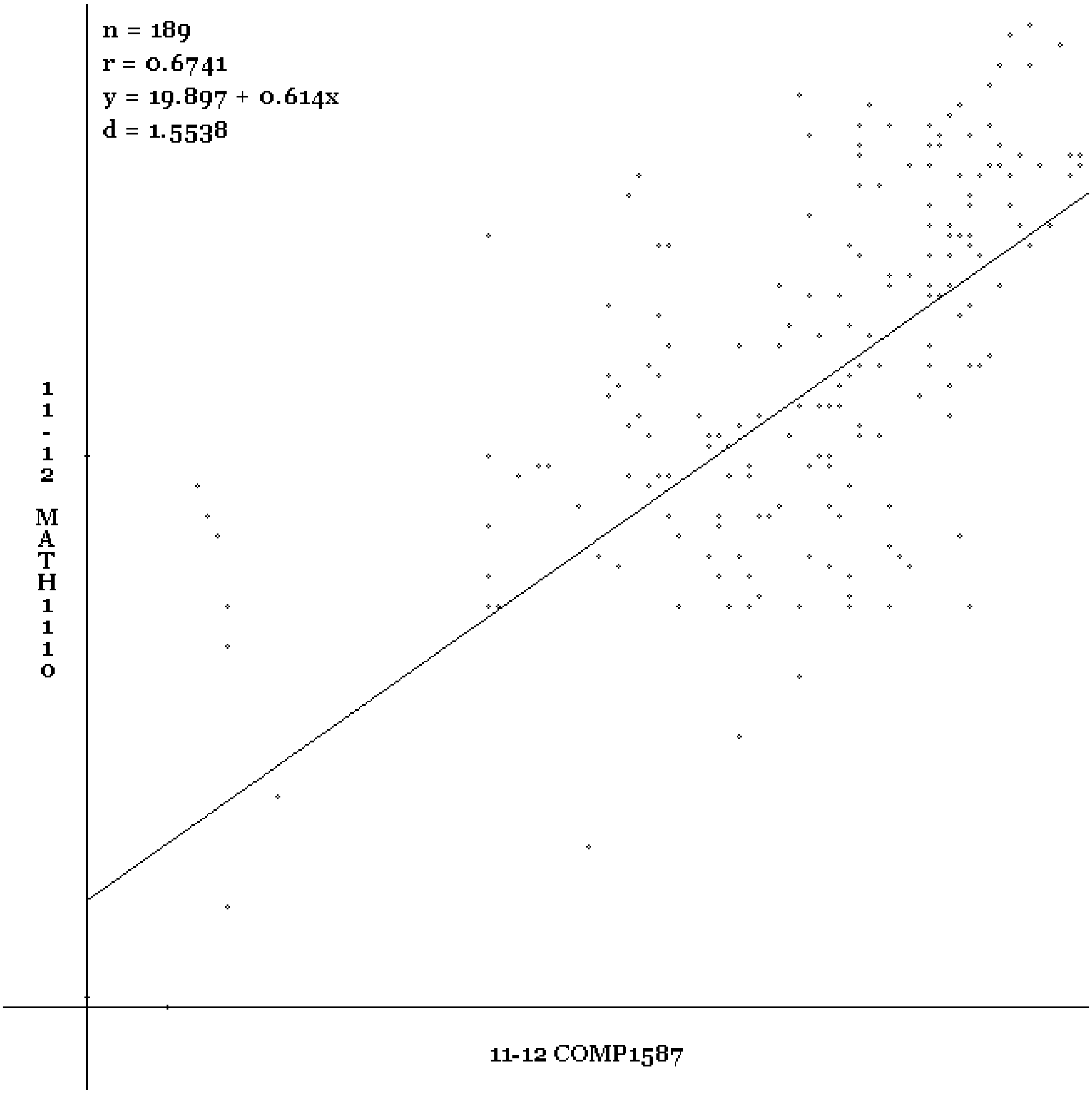


n = 164
r = 0.6203
y = 7.610 + 0.715x
d = 1.6947
11-12 MATH1111
11-12 COMP1587

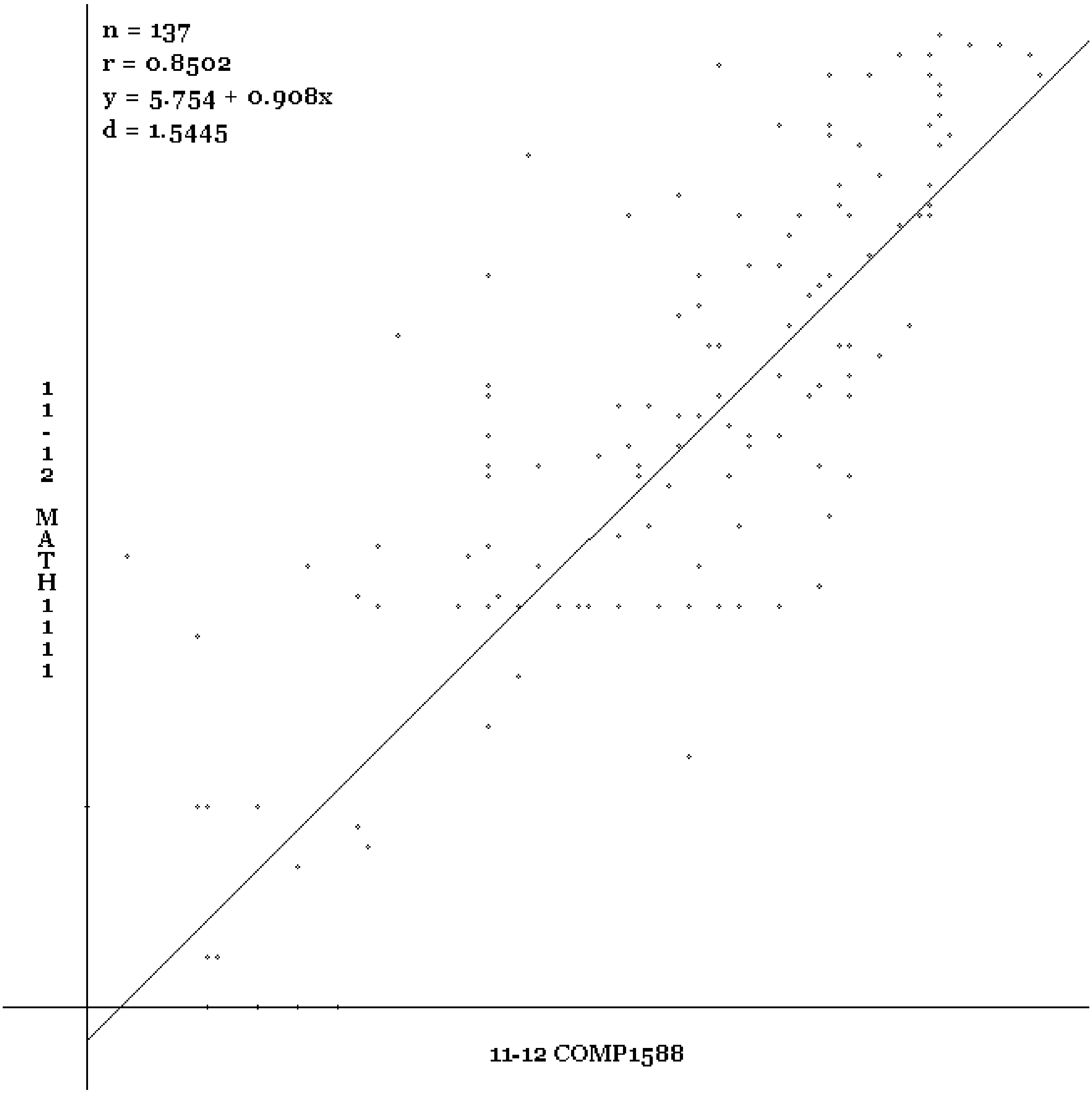


n = 148
r = 0.7046
y = -2.544 + 0.816x
d = 1.7225
11-12 COMP1588
11-12 COMP1587

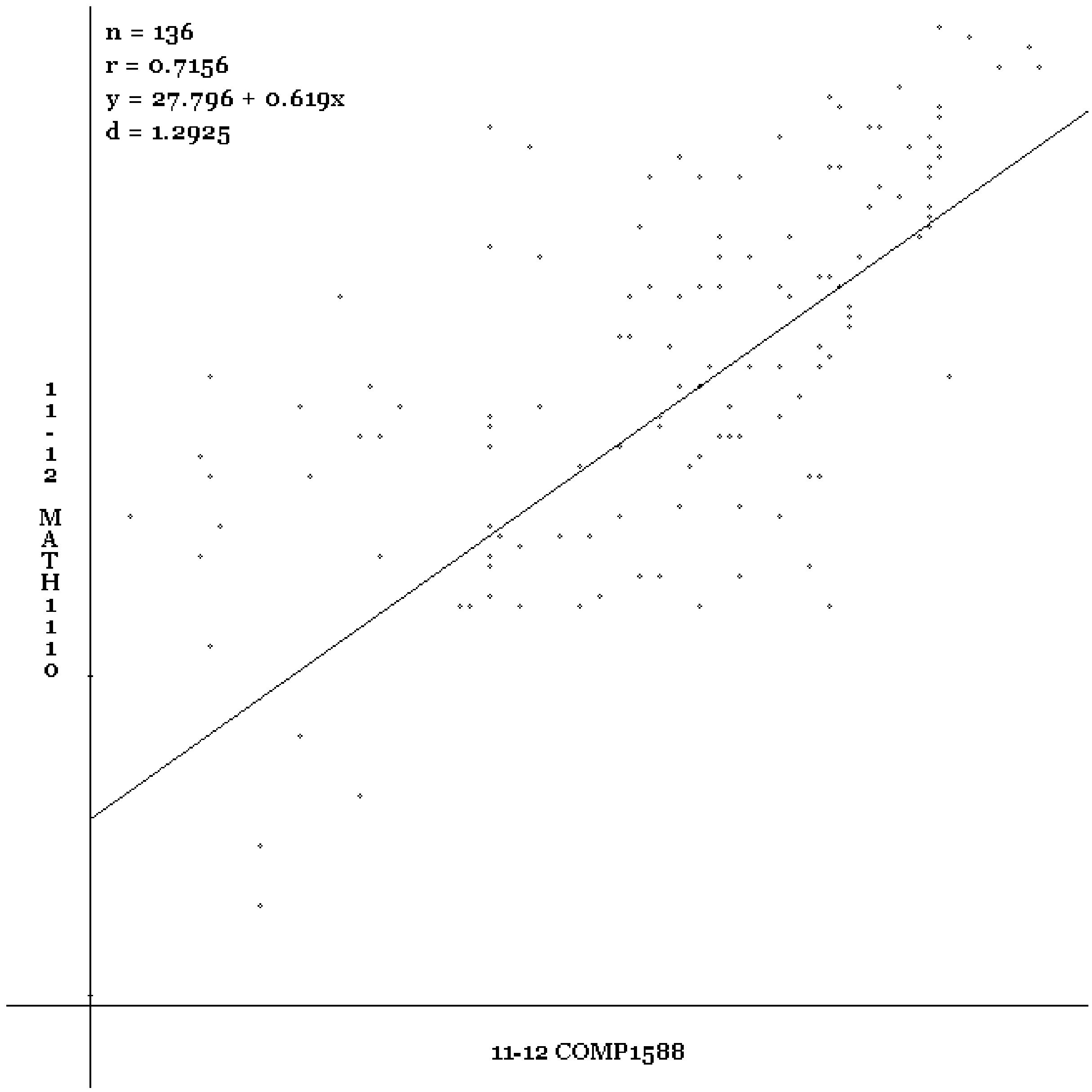

n = 183
r = 0.3966
y = 42.100 + 0.387x
d = 1.5370
11-12 COMP1589
11-12 COMP1587

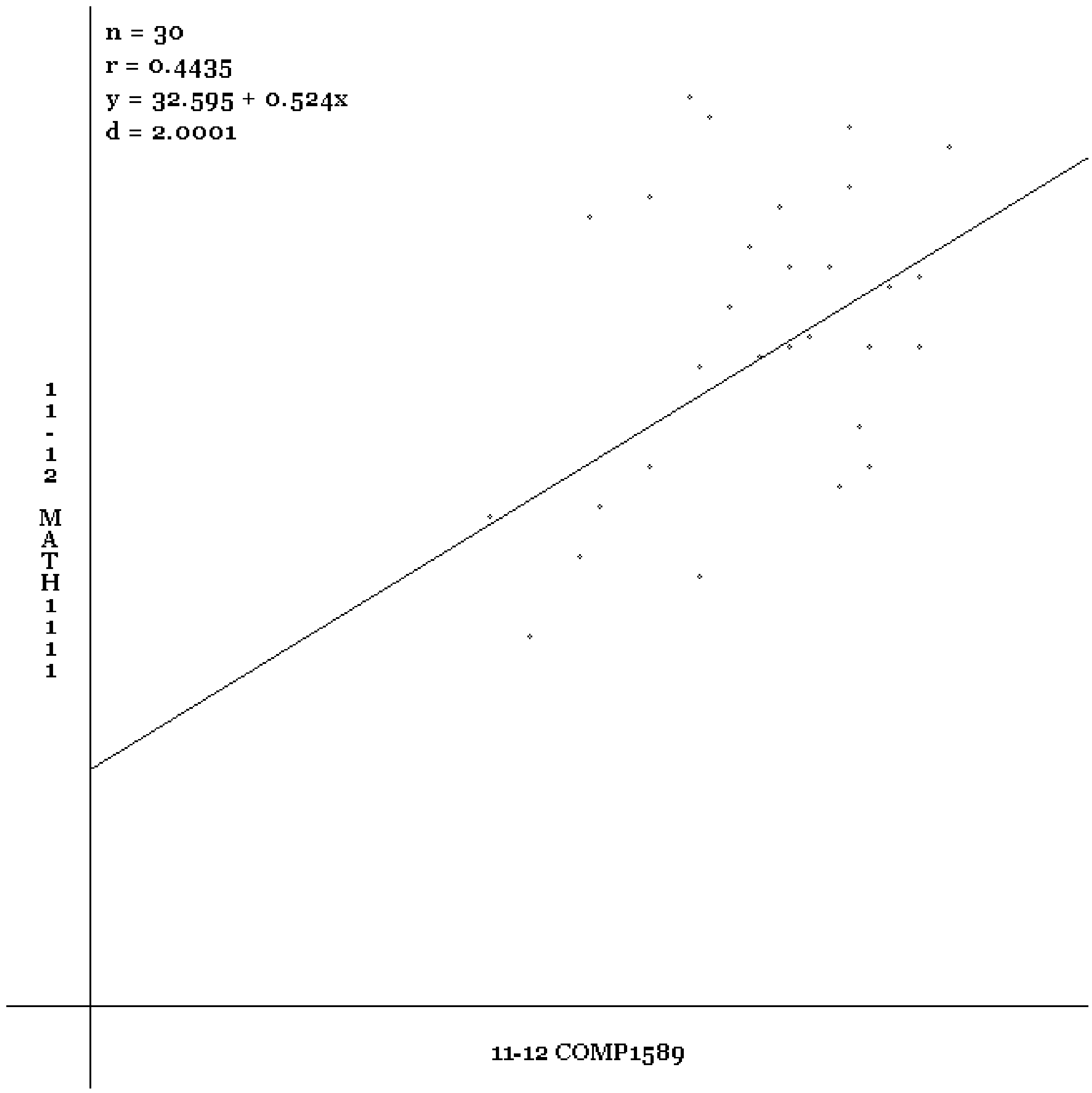


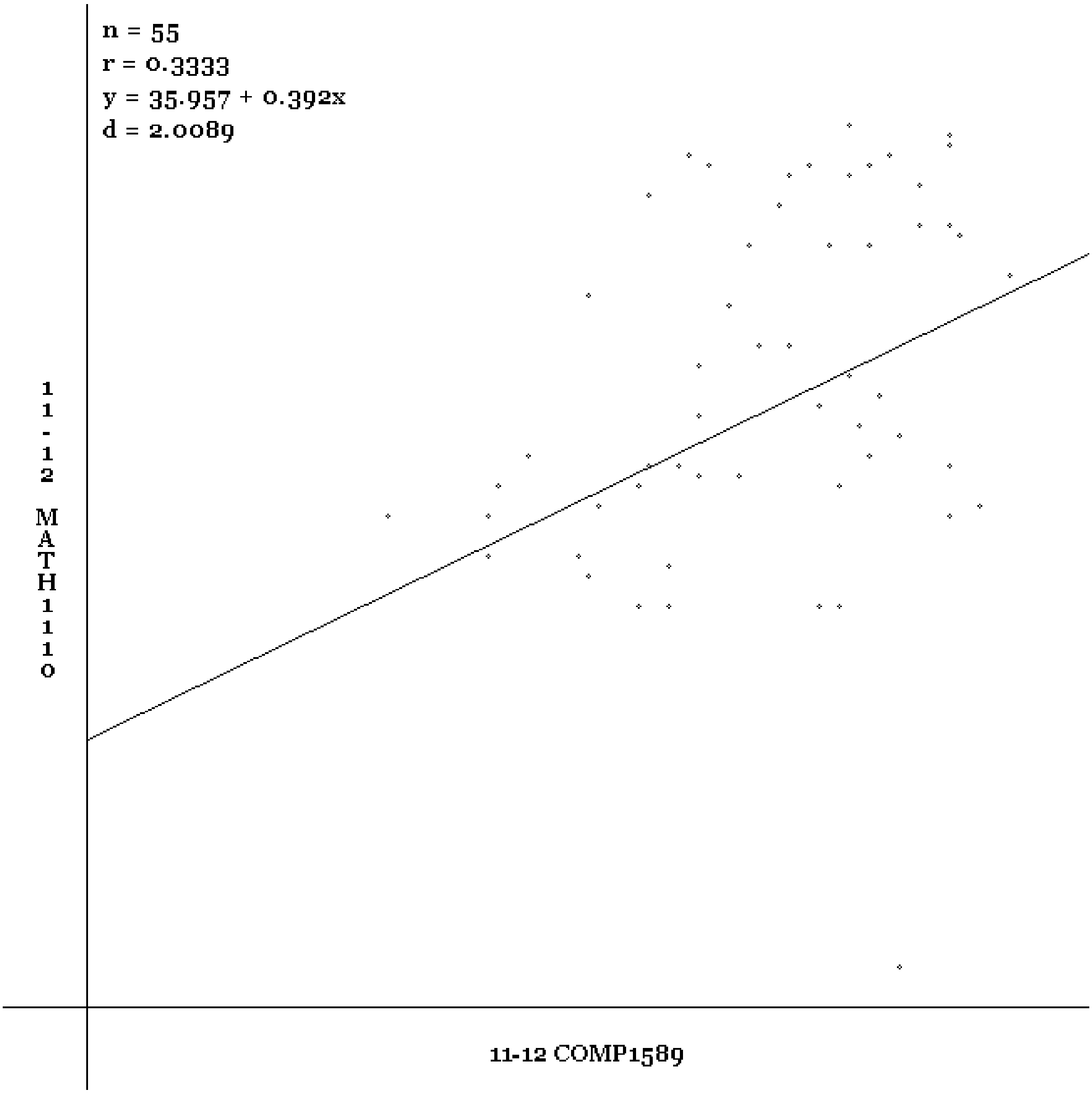
n = 189
r = 0.6741
y = 19.897 + 0.614x
d = 1.5538
11-12 MATH1110
11-12 COMP1587

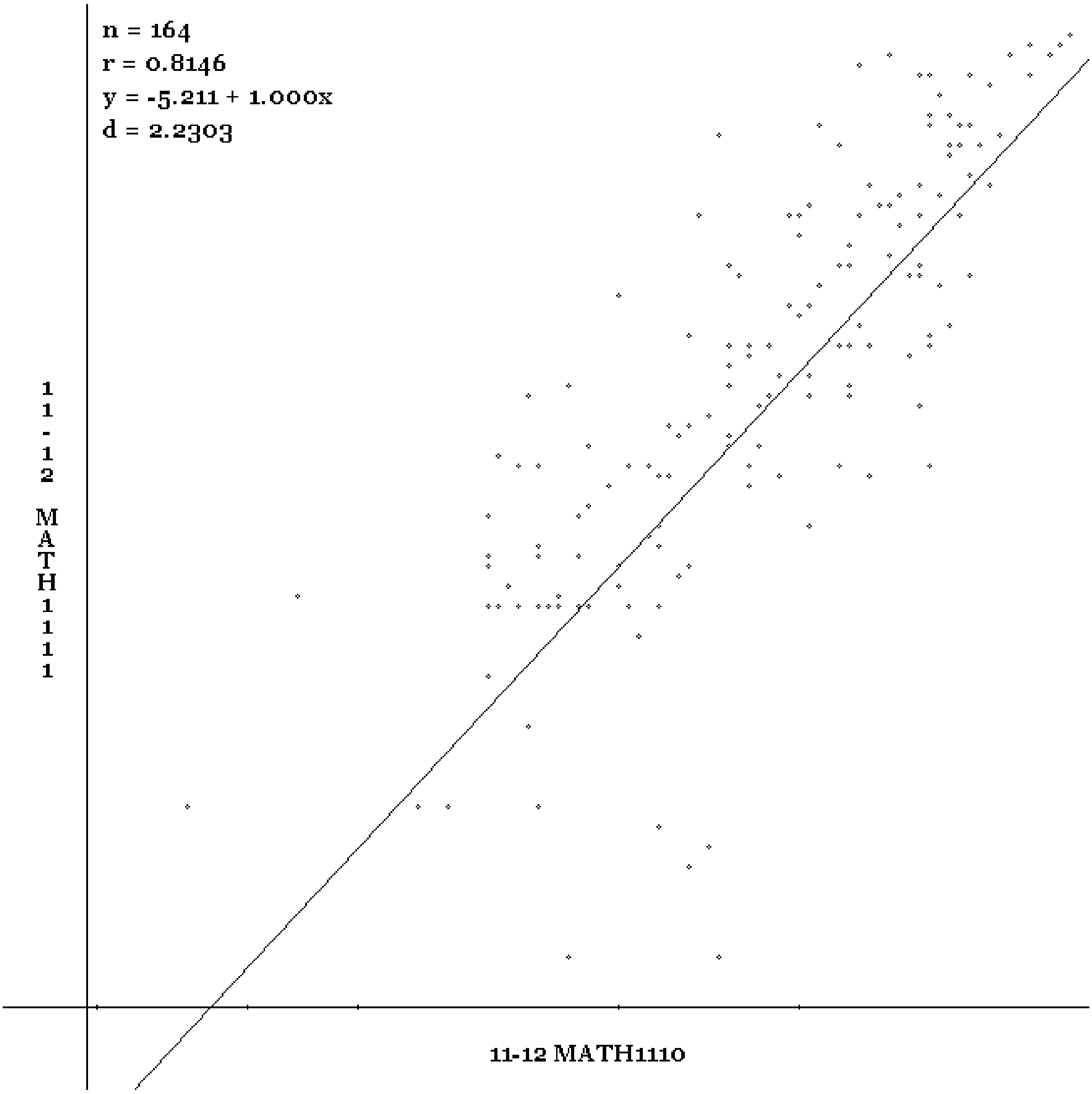


n = 137
r = 0.8502
y = 5.754 + 0.908x
d = 1.5445
11-12 MATH1111
11-12 COMP1588


n = 136
r = 0.7156
y = 27.796 + 0.619x
d = 1.2925
11-12 MATH1110
11-12 COMP1588


n = 30
r = 0.4435
y = 32.595 + 0.524x
d = 2.0001
11-12 MATH1111
11-12 COMP1589


n = 55
r = 0.3333
y = 35.957 + 0.392x
d = 2.0089
11-12 MATH1110
11-12 COMP1589


n = 164
r = 0.8146
y = -5.211 + 1.000x
d = 2.2303
11-12 MATH1111
11-12 MATH1110